FINAL YEAR PROJECT THESIS

Medical Image Processing and Machine Learning on Skin Cancer

Kunjal S. Panchal – 804025

Viral Odedara – 804023

Hardika Mehta - 804049

Janavi Trivedi - 804059

in partial fulfillment for the award of the degree of

Bachelor of Engineering in Computer Science and Engineering

APRIL – 2019

Department of Computer Science and Engineering

Faculty of Technology and Engineering

The Maharaja Sayajirao University of Baroda

सत्यं शिवं सुन्दरम् The Maharaja Sayajirao University of Baroda

Declaration

I hereby declare that the work on the thesis and the project in this thesis is my own except for quotations and summaries which have been duly acknowledged. The thesis has not been accepted for any degree and is not concurrently submitted for award of other degree.

Signature

Kunjal S. Panchal

Seat Number: 804025

PRN Number: 2015033800105024

Date:

Signature

Viral S. Odedara

Seat Number: 804023

PRN Number: 2016033800116023

Date:

Signature

Hardika Mehta

Seat Number: 804049

PRN Number: 2016033800116093

Date:

Signature

Janavi Trivedi

Seat Number: 804059

PRN Number: 2016033800121875

Date:

Place:

Approval/ Certification

This is to certify that the project report entitled Medical Image Processing and Machine Learning on Skin cancer, submitted to the Department of Computer Science and Engineering, The Maharaja Sayajirao University of Baroda, in partial fulfillment for the award of the degree of Bachelor of Engineering in Computer Science and Engineering, is a record of bona fide work carried out by Ms Kunjal Panchal, Seat No. 804025, Ms Viral Odedara, Seat No. 804023, Ms Hardika Mehta, Seat No. 804049 and Ms Janavi Trivedi, Seat No. 804059 under my supervision and guidance.

All help received by them from various sources have been duly acknowledged.

No part of this report has been submitted elsewhere for award of any other degree.

Dr. Mamta Padole

Associate Professor

Supervisor and Guide

Dept. of Computer Science & Engg.

The Maharaja Sayajirao University of Baroda

Date:

Place:

Dr. Apurva Shah

Head of the Department

Dept. of Computer Science & Engg.

The Maharaja Sayajirao University of Baroda

Acknowledgements

This research was supported by The Maharaja Sayajirao University of Baroda. We thank our friends from BE-IV Department of Computer Science and Engineering who provided insight and suggestions that greatly assisted the project.

We thank Dr. Mamta Padole for assistance with Matlab and blueprint of Image Acquisition and Augmentation, and general guidance in every step of the project.

We would also like to show our gratitude to Bharatan Mudaliar and Nisarg Parikh for helping us with Neural Networks and Image Pre-Processing steps.

Abstract

In this thesis, two methods to detect skin cancer from dermatoscopic images of carcinogenic moles are examined: one is of unsupervised learning where we process the image using its visual characteristics and the other one is of supervised learning where we use a neural network to learn decision paths for either the image is malignant or benign and then input any new/ different image to make decision. The image processing part has pre-processing section where we remove hair or scale marks from the image; these pre-processed images will go in both of these method as the input. Few image augmentation techniques for neural network training set are also depicted.

Keywords: Malignant Melanoma, Basal Cell Carcinoma, Squamous Cell Carcinoma, Benign, Skin Cancer, Otsu' Thresholding, Fiber Metrics Filter, Bottom hat Filter, Hair removal, Image Augmentation, MobileNet, Region of Interest clustering, Border irregularity checking, Euclidean Distance, Dice-Sorenson Co-efficient, Histogram Analysis.

Table of Contents

List of Figures:

List of Tables:

List of Equation:

CHAPTER 1

INTRODUCTION

1.1 Introduction

The functionality of skin plays a vital role in the human body since it is the largest organ which covers the muscles, bones and other parts of the body. Once the functionality of skin goes wrong it affects the other parts of the body. Skin is the most sensitive part, therefore when it is explored into sunlight and other environmental pollution tends to occur skin cancer. Skin cancer ap-pears to be of two kinds Benign and Melanoma form. Benign it's just the moles on the skin which does not penetrate inside, where-as Melanoma causes sores on the skin which leads to bleeding and it is named after cells Melanocytes which is more hazardous. In United States, more than 700,000 skin lesions are diagnosed annually under the estimation of American Cancer Society. According to statistics given by the Apollo and other hospitals it suggests that Melanoma affects the ages ranging from 41-60+. There are technologies that are used to detect skin cancer at the early stages. Skin Cancer detected in advance can save people's lives and it eliminates the multiplication of cancer cells across the parts of the body. Although it affects the people within age limits but high probably is for the bright skin people. [1]

1.2 Need for awareness of Skin Cancer

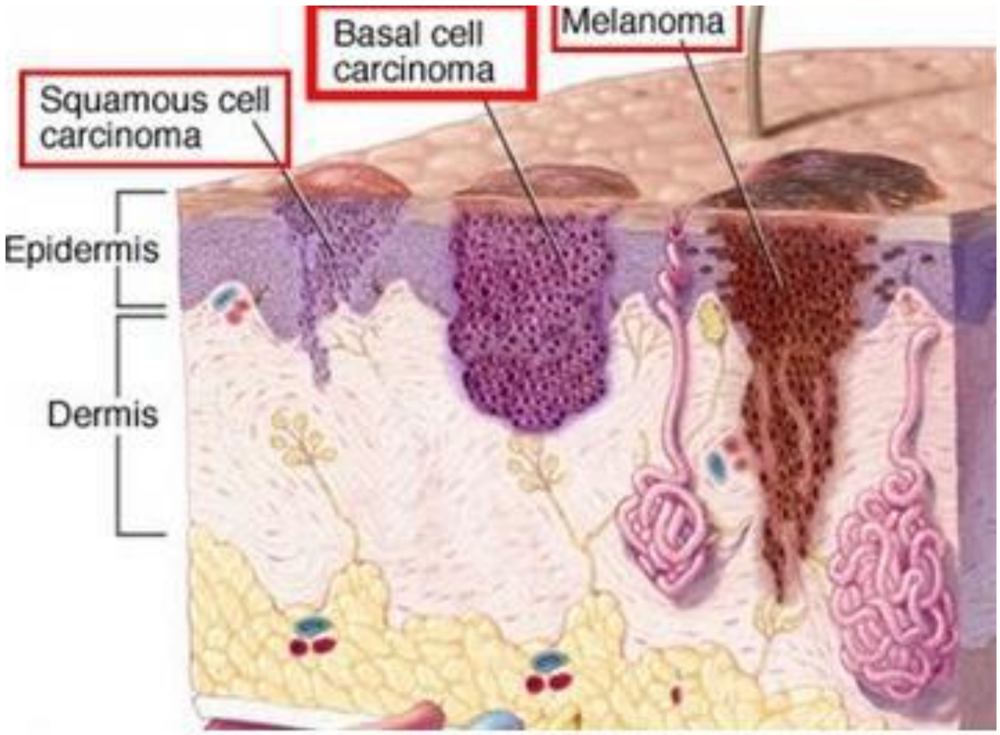


Fig 1: states the different types of skin cancer; the diagram clearly shows the Melanoma penetration inside the skin.

It will be hard for even the experienced dermatologist to detect the skin cancer or to predict the stages. Therefore, many hardware & software devices and applications evolved. In order to detect Non- Melanoma Skin Cancer (NMSC) there are many hardware devices which includes, biopsy, molecular markers, ultra-sonography, Doppler, optical coherence tomography, Dermoscopy and Spectroscopy. Dermoscopy is one such hardware device that helps in examining the surface of the skin using skin surface microscopy for the detection of skin cancer; it is a non-invasive examination technique that lends the support to distinguish between Melanoma and Non-Melanoma. It is the most expensive device that cannot be affordable for the poor countries and developing countries like Africa. Moreover, later researchers came up with many software applications, which eliminated the hardware cost and influence the ease of usage. The software does not require any expertise to handle them but developed with the intention of prior information about the cancer. Image Processing is one of the traditional approaches which deals with analyzing and processing of an image. Image processing has proposed many methodologies which helps in the early detection of skin cancer. Proposed methodologies have influenced in the early detection which prevents the cancer from spreading across the skin. The methodology in an image processing approach involves the use of Noise removal, Edge detection, Image enhancement, Segmentation algorithms, Feature selection, Feature extraction, calculating Area, Perimeter, Eccentricity, Neural network approach using Back propagation algorithms. In this paper, a model is proposed with an aim of curing the skin cancer along with providing cancer information to the people. Model detects the skin cancer on the color skin image and uses pre-processing methods such as Image acquisition, Noise removal and plotting Histogram. Pre-processing methods helps to solve the illumination, contrast and noise problems. After the removal of noise various techniques such as Edge detection, Image enhancement, Segmentation. Feature extraction is used to extract the affected portion of skin, then the Area, Perimeter and Eccentricity is calculated. Calculated values are fed into the Neural networks, using Back propagation algorithm the stage and type of the skin cancer is predicted. The main aim of the proposed system is to eliminate the risk caused by many countries with respect to the skin cancer. Many hardware devices were developed but those devices were not affordable. Patients are not only provided with the information of type of cancer but appropriate stage of the skin cancer is predicted which helps to easily cure the cancer when stage is obtained. [1]

1.3 Types of Skin Cancer

There are three main types of skin cancer: basal-cell skin cancer (basal-cell carcinoma) (BCC), squamous-cell skin cancer (squamous-cell carcinoma) (SCC) and malignant melanoma. Less common skin cancers include: dermatofibrosarcoma protuberans, Merkel cell carcinoma, Kaposi's sarcoma, keratoacanthoma, spindle cell tumors, sebaceous carcinomas, microcystic adnexal carcinoma, Paget's disease of the breast, atypical fibroxanthoma, leiomyosarcoma, and angiosarcoma. We will put these less common types plus all moles, acne, freckles into “benign” category.

1.3.1 Basal Cell Carcinoma

Basal-cell carcinomas are present on sun-exposed areas of the skin, especially the face. They rarely metastasize and rarely cause death. They are easily treated with surgery or radiation.

Basal-cell skin cancer (BCC) usually presents as a raised, smooth, pearly bump on the sun-exposed skin of the head, neck or shoulders.

Sometimes small blood vessels (called telangiectasia) can be seen within the tumor. Crusting and bleeding in the center of the tumor frequently develops. It is often mistaken for a sore that does not heal. This form of skin cancer is the least deadly and with proper treatment can be completely eliminated, often without scarring.

1.3.2 Squamous Cell Carcinoma

Squamous-cell skin cancer are common, but much less common than basal-cell cancers. They metastasize more frequently than BCCs. Even then, the metastasis rate is quite low, with the exception of SCC of the lip, ear, and in people who are immunosuppressed.

Squamous-cell skin cancer are common, but much less common than basal-cell cancers. They metastasize more frequently than BCCs. Even then, the metastasis rate is quite low, with the exception of SCC of the lip, ear, and in people who are immunosuppressed.

1.3.3 Malignant Melanoma

Melanoma are the least frequent of the 3 common skin cancers. They frequently metastasize, and could potentially cause death once they spread.

Most melanoma consist of various colors from shades of brown to black. A small number of melanomas are pink, red or fleshy in color; these are called amelanotic melanoma and tend to be more aggressive. Warning signs of malignant melanoma include change in the size, shape, color or elevation of a mole. Other signs are the appearance of a new mole during adulthood or pain, itching, ulceration, redness around the site, or bleeding at the site. An often-used mnemonic is "ABCDE", where A is for "asymmetrical", B for "borders" (irregular: "Coast of Maine sign"), C for "color" (variegated), D for "diameter" (larger than 6 mm – the size of a pencil eraser) and E for "evolving. [8]

1.4 Signs and Symptoms

Basal Cell Carcinoma:

Fig 2: Basal Cell Carcinoma

Note the pearly translucency to fleshy color, tiny blood vessels on the surface, and sometime ulceration which can be characteristics. The key term is translucency.

Basal cell cancers usually develop on areas exposed to the sun, especially the face, head, and neck, but they can occur anywhere on the body.

These cancers can appear as:

- Flat, firm, pale or yellow areas, similar to a scar
- Raised reddish patches that might be itchy
- Small, pink or red, translucent, shiny, pearly bumps, which might have blue, brown, or black areas
- Pink growths with raised edges and a lower area in their center, which might contain abnormal blood vessels spreading out like the spokes of a wheel
- Open sores (which may have oozing or crusted areas) that don't heal, or that heal and then come back
- Basal cell cancers are often fragile and might bleed after shaving or after a minor injury. Sometimes people go to the doctor because they have a sore or a cut from shaving that just won't heal, which turns out to be a basal cell cancer. A simple rule of thumb is that most shaving cuts heal within a week or so.

Squamous Cell Carcinoma:

Fig 3: Squamous Cell Carcinoma:

Commonly presents as a red, crusted, or scaly patch or bump. Often a very rapid growing tumor. Squamous cell cancers tend to occur on sun-exposed areas of the body such as the face, ear, neck, lip, and back of the hands. Less often, they form in the skin of the genital area. They can also develop in scars or skin sores elsewhere.

These cancers can appear as:

- Rough or scaly red patches, which might crust or bleed
- Raised growths or lumps, sometimes with a lower area in the center
- Open sores (which may have oozing or crusted areas) that don't heal, or that heal and then come back
- Wart-like growths

Malignant Melanoma:

The ABCDE rule is another guide to the usual signs of melanoma. Be on the lookout and tell your doctor about spots that have any of the following features:

- A is for Asymmetry: One half of a mole or birthmark does not match the other.
- B is for Border: The edges are irregular, ragged, notched, or blurred.
- C is for Color: The color is not the same all over and may include different shades of brown or black, or sometimes with patches of pink, red, white, or blue.
- D is for Diameter: The spot is larger than 6 millimeters across (about ¼ inch – the size of a pencil eraser), although melanomas can sometimes be smaller than this.
- E is for Evolving: The mole is changing in size, shape, or color.

Normal Moles:

A normal mole is usually an evenly colored brown, tan, or black spot on the skin. It can be either flat or raised. It can be round or oval. Moles are generally less than 6 millimeters (about ¼ inch) across (about the width of a pencil eraser). Some moles can be present at birth, but most appear during childhood or young adulthood. New moles that appear later in life should be checked by a doctor. [8]

Once a mole has developed, it will usually stay the same size, shape, and color for many years. Some moles may eventually fade away.

Most people have moles, and almost all moles are harmless. But it's important to recognize changes in a mole – such as in its size, shape, or color – that can suggest a melanoma may be developing. [7, 8]

CHAPTER 2

BASICS OF DIGITAL IMAGE PROCESSING

2.1Filters

Image filtering is useful for many applications, including smoothing, sharpening, removing noise, and edge detection. A filter is defined by a kernel, which is a small array applied to each pixel and its neighbors within an image. In most applications, the center of the kernel is aligned with the current pixel, and is a square with an odd number (3, 5, 7, etc.) of elements in each dimension. The process used to apply filters to an image is known as convolution, and may be applied in either the spatial or frequency domain.

2.1.1 Median Filter

Median filtering is a nonlinear operation often used in image processing to reduce "salt and pepper" noise. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges.

While performing median filtering of the matrix A in two dimensions; each output pixel contains the median value in a 3-by-3 neighborhood around the corresponding pixel in the input image. medfilt2 pads the image with 0's on the edges, so the median values for points within one-half the width of the neighborhood ([m n]/2) of the edges might appear distorted.

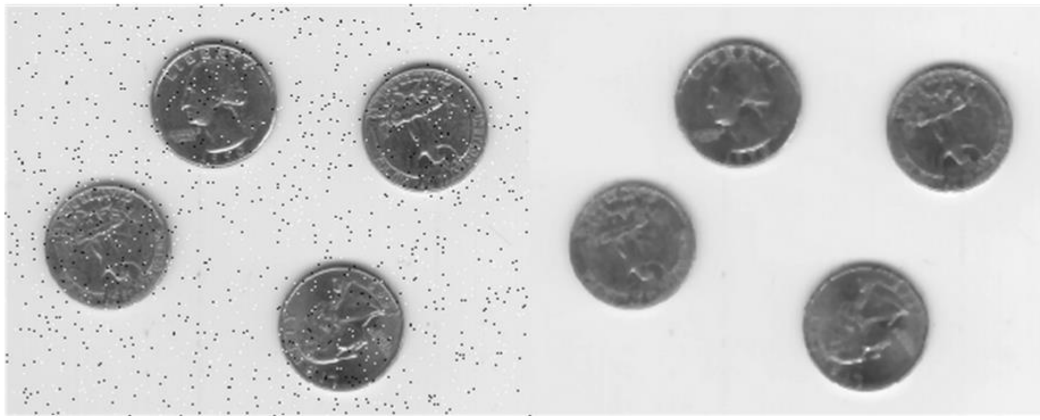


Fig 4: Before and after applying median filtering

2.1.2 Bottom Hat Filter

A strel object represents a flat morphological structuring element, which is an essential part of morphological dilation and erosion operations.

A flat structuring element is a binary valued neighborhood, either 2-D or multidimensional, in which the true pixels are included in the morphological computation, and the false pixels are not. The center pixel of the structuring element, called the origin, identifies the pixel in the image being processed. Use the strel function to create a flat structuring element. You can use flat structuring elements with both binary and grayscale images. The following figure illustrates a flat structuring element.

Bottom Hat Filter performs morphological bottom-hat filtering on the grayscale or binary input image, IM, returning the filtered image, IM2. SE is a structuring element returned by the strel function. SE must be a single structuring element object, not an array containing multiple structuring element objects.

2.1.3 Fiber Metrics Filter

For enhance elongated or tubular structures in image.

FiberMetric Filter enhances elongated or tubular structures in intensity image A using Hessian-based multiscale filtering. The image returned, B, contains the maximum response of the filter at a thickness that approximately matches the size of the tubular structure in the image.

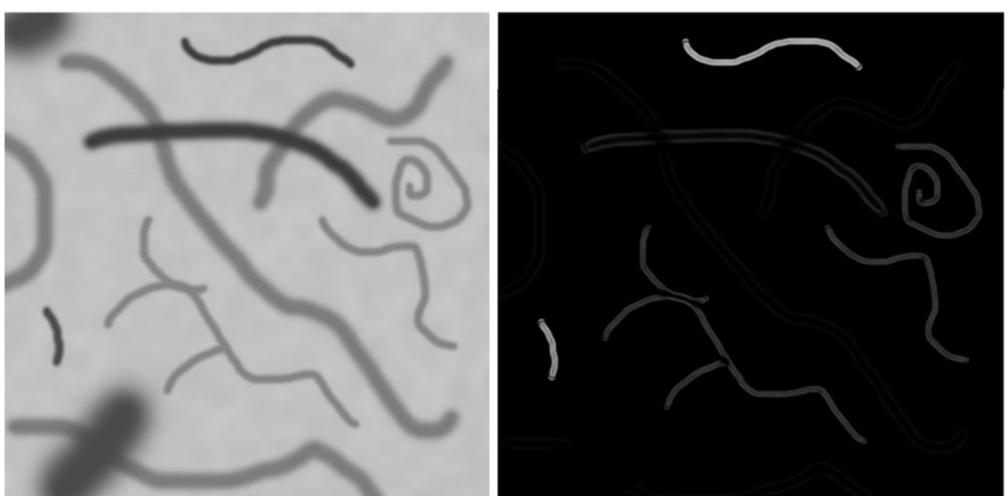


Fig 5: Before and after applying fiber metrics filtering

The filter enhances an image to highlight structures and is typically used as a preprocessing step for segmentation.

2.1.4 Black and White Region Fill

Remove small objects from binary image.

It removes all connected components (objects) that have fewer than P pixels from the binary image BW, producing another binary image, BW2. The default connectivity is 8 for two dimensions, 26 for three dimensions, and conndef(ndims(BW), 'maximal') for higher dimensions. This operation is known as an area opening.

2.2 Region Filling and Interpolation

Fill in specified regions in image using inward interpolation.

It fills the regions in image I specified by mask. Non-zero pixels in mask designate the pixels of image I to fill. You can use regionfill to remove objects in an image or to replace invalid pixel values using their neighbors.

Specify a polygon that completely surrounds one of the coins in the image. This example uses the x-coordinates and y-coordinates (columns and rows) of the polygon vertices to specify the region.

It smoothly interpolates inward from the pixel values on the outer boundary of the regions. It computes the discrete Laplacian over the regions and solves the Dirichlet boundary value problem.

2.3 Thresholding

Convert image to binary image, based on threshold "**level**".

It converts the grayscale image I to a binary image. The output image BW replaces all pixels in the input image with luminance greater than level with the value 1 (white) and replaces all other pixels with the value 0 (black). Specify level in the range [0,1]. This range is relative to the signal levels possible for the image's class. Therefore, a level value of 0.5 is midway between black and white, regardless of class. To compute the level argument, you can use the function graythresh. If you do not specify level, im2bw uses the value 0.5.

Fig 6: Before and after thresholding a RGB image

2.4 Contrast Adjustment and Sharpening

Adjust image intensity values or colormap.

It maps the intensity values in grayscale image I to new values in J such that 1% of data is saturated at low and high intensities of I. This increases the contrast of the output image J.

Contrast limits for input image, specified as a two-element numeric vector with values between 0 and 1. Values below low_in and above high_in are clipped; that is, values below low_in map to low_out, and those above high_in map to high_out. If you specify an empty matrix ([]), imadjust uses the default limits [0 1].

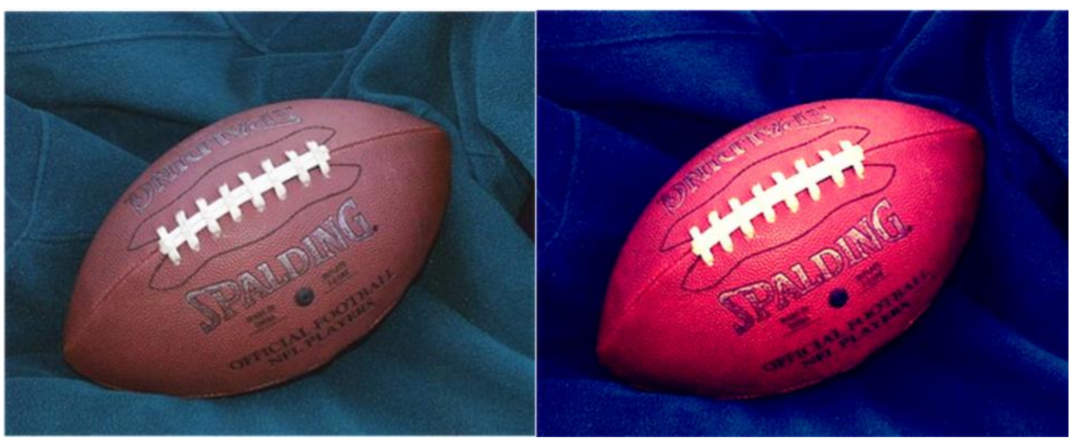


Fig 7: Before and after adjusting contrast of a RGB image

Grayscale colormap for contrast enhancement: enhances the contrast of an image. It creates a new gray colormap, cmap, that has an approximately equal intensity distribution. All three elements in each row are identical. It returns a gray colormap that is the same length as the current colormap. If there are NaN or Inf elements in X the length of the colormap increases.

2.5 RGB Channels

White light (sunlight) contains all the colors that humans can see. This can be demonstrated by channeling light through a prism, which results in a rainbow. The same thing can be seen in nature, when sunlight is shining through drops of water. White light can also be constructed by taking red, green, and blue light and pointing them on top of each other in a dark room, like is shown in the figure. It can be seen that the color in the center will be white.

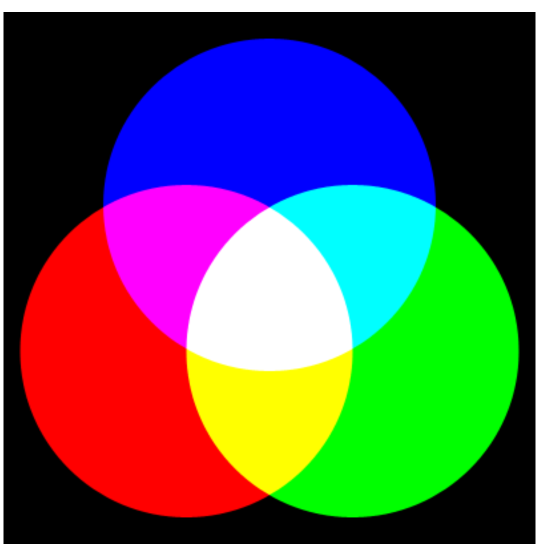


Fig 8: RGB color system

More variations can be created by changing the color of one of the spotlights. This same phenomenon is used in computer graphics in an RGB color system. The center color changes when one of the light components changes brightness. In the RGB color space, each of these color values are represented as a value between 0 and 255. "RGB" stands for "red", "green" and "blue", each of which point to the value of the specific color component. White color can be formed when each of these components are 255. A bright red color can be formed by setting the red value to 255 and the other values to 0 and so on. In computer graphics, the color values can also be represented as values between 0 and 1, so they can be used in vector multiplication. The color can be transformed to the 0 to 1 system by dividing the color component values by 255. RGB system has 3 separate components, so it can be stored in a 3D vector and use in vector and matrix calculations. The RGB system can also be expanded to the RGBA system, which has an extra component presenting the alpha channel (transparency).

2.6 Segmentation

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The simplest method of image segmentation is called the thresholding method. This method is based on a clip-level (or a threshold value) to turn a gray-scale image into a binary image.

The key of this method is to select the threshold value (or values when multiple-levels are selected). Several popular methods are used in industry including the maximum entropy method, balanced histogram thresholding, Otsu's method (maximum variance), and k-means clustering.

Recently, methods have been developed for thresholding computed tomography (CT) images. The key idea is that, unlike Otsu's method, the thresholds are derived from the radiographs instead of the (reconstructed) image.

New methods suggested the usage of multi-dimensional fuzzy rule-based non-linear thresholds. In these works decision over each pixel's membership to a segment is based on multi-dimensional rules derived from fuzzy logic and evolutionary algorithms based on image lighting environment and application.

CHAPTER 3

PRE-PROCESSING

3.1Need for pre-processing

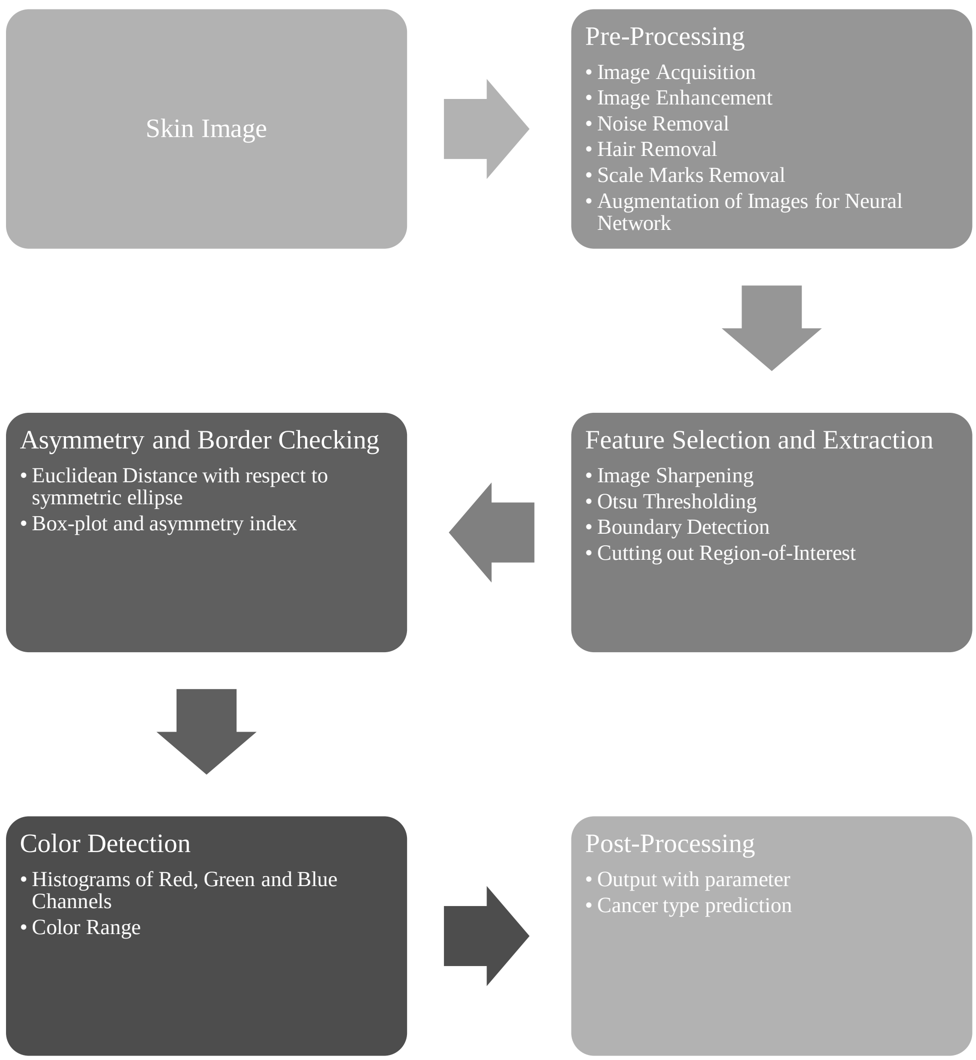


Fig 9: Blueprint for Image Processing to detect subtype of skin cancer

Image pre-processing is an essential step of detection in order to remove noises and enhance the quality of original image. It required to be applied to limit the search of abnormalities in the background influence on the result. The main purpose of this step is to improve the quality of melanoma image by removing unrelated and surplus parts in the back ground of image for further processing. Good selection of preprocessing techniques can greatly improve the accuracy of the system. [3]

Image Enhancement is a crucial procedure to improve the visual appearance of the image; it is defined as provider of the "better" transform representation for further automated steps of detection. Thus, the image enhancement can be categorized in three categories:

Image Scaling - Image scaling techniques are applied due to the lack of same and standard size of images. Since the skin cancer images may be gathered from different sources and sizes, the first step is to resize the images to have the fixed width pixels but variable size of height.

Color Space Transformation- Since color information plays an inevitable role in skin cancer detection systems, researchers try to extract the more corresponding color of images for further processing. Generally, the common color spaces include RGB, HSV, HSI, CIE-LAB and CIE-XYZ. RGB is a color space which comprise the red, green, and blue spectral wavelength. The most frequently presentation of colors in image processing is RGB. Since RGB color space has some limitation in high level processing, other color space representations have been developed. HSV and HSI color spaces imitate the human visual perception of color in terms of hue, saturation and intensity which are respectively the average wavelength of the color, the amount of white in the color and the brightness. The next color space is CIE-LAB which has been proposed to provide uniformity. CIE-XYZ is another color space which can produce every color with positive tristimulus values. Since the purpose in images of skin cancer detection systems is to obtain the high-level variations between intensities to detect the edges of lesions, it would be optimal to convert the image into gray scale. Since LAB is one of the useful color models which represent every color through three components of luminance, red/green and blue/yellow, it could be beneficial to transform the RGB to LAB using XYZ as an intermediate colorspace.

Contrast enhancement is beneficial step to improve the perception for further processing; it can sharpen the image border and improve the accuracy by accentuating the brightness difference between background and foreground. Contrast enhancement plays a vital role in increasing the quality of an image. The widely practiced methods are classified into “Linear contrast enhancement” and “Non-Linear contrast enhancement” techniques. Linear contrast enhancement mostly used in remotely sensed images.

3.2 Image Acquisition

Image analysis starts with image acquisition this involves all aspects that have to be addressed in order to obtain dermoscopy image of human skin cancer the selection of radiation (light) sources and sensors (such as cameras) has to be considered very carefully. For this study, images have been taken from the following websites:

https://www.isic-archive.com/

https://www.dermnetnz.org/topics/dermoscopy/

http://www.fc.up.pt/addi/ph2%20database.html

https://www.derm101.com/image-library/?match=IN

3.3 Hair and Scale Marks Removal

Although the thin blood vessels and skin lines will be smoothed using most of restoration filters, the image may include the thick hairs .Thick hairs in automated analysis of small skin lesions are considered as a common impediment which are able to mislead the segmentation process. To remove the thick hairs in skin cancer images, researchers applied other methods such as mathematical morphology methods, curvilinear structure detection, an inpainting based method approach, automated software called DullRazor and Top Hat transform combined with a bicubic interpolation approach. The hair-free images are acquired using the operations. [4]

3.3.1 Extracting Color Channels

To apply the rest of the process we have designed; we need to get all 3 RGB layers of the image separately. We will acquire Red, Green and Blue Channels of the Original colored skin image.

3.3.2 Applying Filters

Our focus is to get hair-shaped structures extracted. For this, the best technique is to use morphological structural elements. A flat morphological structuring element, which is an essential part of morphological dilation and erosion operations.

A flat structuring element is a binary valued neighborhood, either 2-D or multidimensional, in which the true pixels are included in the morphological computation, and the false pixels are not. The center pixel of the structuring element, called the origin, identifies the pixel in the image being processed. The following figure illustrates a flat structuring element.

The shape we went for creates a disk-shaped structuring element, where r specifies the radius. n specifies the number of line structuring elements used to approximate the disk shape. Morphological operations using disk approximations run much faster when the structuring element uses approximations.

3.3.3 Acquiring the Binary Mask

After getting the grayscale image of hair-like structure, we convert it to binary by a pre-defined threshold level to get the mask which will be used to determine which pixel values are to be interpolated.

3.3.4 Region Filling the Rest of the Image

The original RGB skin image will be hypothetically superposed by the binary mask of hair-like structure and the values which represent hair pixels will be changed to interpolated approximations calculated from its non-hair neighboring pixel. The pixel values will be "filled" according to neighboring values, giving it almost same pattern, texture and color.

3.4 Augmenting Images for Neural Network Training

3.4.1 Rotation

Rotations in computer graphics is a transformational operation. That means that it is a conversion from one coordinate space onto another. Rotational transformation can be accomplished with Matrices or with Quaternions. You will learn how a vector can be rotated with both methods. Although Quaternions offer a better solution than matrices, it is a good idea to learn how matrices rotate a character in 3D games.

The matrix for rotating a point about an origin in a 2D plane is defined as:

$$R_{\beta}=\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$

Equation 1: Rotation Matrix to Roate by β angle

Thus, the rotation of a 2D vector in a plane is done as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix}=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

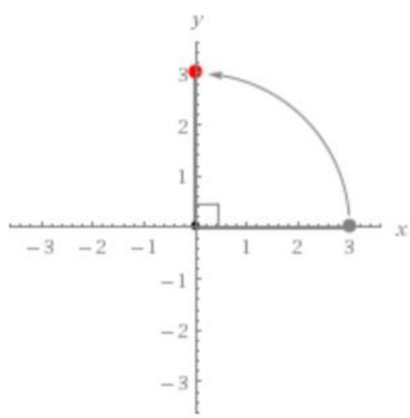
Equation 2: Rotate Each Image Pixel by θ angle

For example: To rotate a vector 90 degrees counterclock-wise is done as follows:

$$\vec{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$



Equation 3: Example Rotation

If you would like to rotate a point about the x-axis, the x-coordinate is kept constant while the y-and z-coordinate are changed as shown below:

$$R_{\beta,x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix}$$

Equation 4: Rotation by β angle around only X-axis

In the same manner, to rotate about the y-axis, the y-coordinate remains constant, while the x- and z-coordinates are changed as shown below.

$$R_{\beta,y}=\begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$$

Equation 5: Rotation by β angle around only Y-axis

3.4.2 Mirroring

A flipped image or reversed image, the more formal term, is a static or moving image that is generated by a mirror-reversal of an original across a horizontal axis (a flopped image is mirrored across the vertical axis). Many printmaking techniques produce images where the printed copy is reversed from the image made on the printing plate, so in a print copying another image, or a real scene or object, unless the artist deliberately creates the plate as a mirror-image of his subject, the finished print will be a mirror image of it. Many print makers developed the skill of reversing images when making printing plates, but many prints, especially early ones, have images that are reversed.

3.4.3 Changing the tint

The RGB characteristics play important role in Neural Network decisions. Thus, we can change values of Red, Green and/or Blue channels by +/- 10~20 to give it a slight change in color, which won't skew the results drastically but still give a different image.

Dermatoscope lenses can be of different tints too, so training dataset having tinted images will enhance the variety of possible real-life images a Neural Network can encounter in its testing or application phase.

CHAPTER 4

IMAGE PROCESSING

4.1 Finding Region of Interest

A region of interest (often abbreviated ROI), are samples within a data set identified for a particular purpose. The concept of a ROI is commonly used in many application areas. For example, in medical imaging, the boundaries of a tumor may be defined on an image or in a volume, for the purpose of measuring its size. The endocardial border may be defined on an image, perhaps during different phases of the cardiac cycle, for example, end-systole and end-diastole, for the purpose of assessing cardiac function. In geographical information systems (GIS), a ROI can be taken literally as a polygonal selection from a 2D map. In computer vision and optical character recognition, the ROI defines the borders of an object under consideration. In many applications, symbolic (textual) labels are added to a ROI, to describe its content in a compact manner. Within a ROI may lie individual points of interest (POIs).

Filtering a region of interest (ROI) is the process of applying a filter to a region in an image, where a binary mask defines the region. For example, you can apply an intensity adjustment filter to certain regions of an image. [5]

To filter an ROI in an image, use the roifilt2 function. You specify:

- Input grayscale image to be filtered
- Binary mask image that defines the ROI
- Filter (either a 2-D filter or function)

The binary mask defines a region of interest (ROI) of the original image. Mask pixel values of 1 indicate the image pixel belongs to the ROI. Mask pixel values of 0 indicate the image pixel is part of the background.

Any binary image can be used as a mask, provided that the binary image is the same size as the image being filtered.

You can create a mask from a grayscale image by classifying each pixel as belonging to either the region of interest or the background. For example, suppose you want to filter the grayscale image I, filtering only those pixels whose values are greater than 0.5.

4.2 Finding Border

After getting ROI extracted, there will be still some skin part which we need to remove for latter stages of color detection and the current one of border detection. For that, we first use Otsu Thresholding technique.

In computer vision and image processing, Otsu's method, named after Nobuyuki Otsu, is used to automatically perform clustering-based image thresholding, or, the reduction of a graylevel image to a binary image. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal.

Fig 10: Applying Otsu's method to convert grayscale image to binary and clustering the region of interest

Now, the ROI will be in logical black and rest of the skin part will be in logical white. This makes finding border of ROI relatively easier. What can be done is as follows:

We scan each row and column of the image matrix. Whenever we get a white pixel after a black one or black pixel after a white one, we understand that the image has been transitioned from background to foreground or vice versa, respectively. This two-nested loop overhead is for getting precise borders for the next stage.

4.3 Check Border for Asymmetry and Irregularity

Asymmetry is the absence of, or a violation of, symmetry (the property of an object being invariant to a transformation, such as reflection). Symmetry is an important property of both physical and abstract systems and it may be displayed in precise terms or in more aesthetic terms. The absence of or violation of symmetry that are either expected or desired can have important consequences for a system.

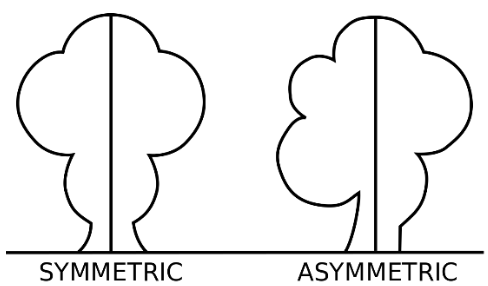


Fig 11: Symmetric versus Asymmetric shape

4.3.1 Euclidean Distance

In some cases, the data are normalized before applying distance calculations. This involves transforming the data to fall within a smaller or common range, such as [1, 1] or [0.0, 1.0]. Consider a height attribute, for example, which could be measured in either meters or inches. In general, expressing an attribute in smaller units will lead to a larger range for that attribute, and thus tend to give such attributes greater effect or "weight." Normalizing the data attempts to give all attributes an equal weight. It may or may not be useful in a particular application. The most popular distance measure is Euclidean distance (i.e., straight line or "as the crow flies"). Let i = (xi1, xi2, . . ., xip) and j = (xj1, xj2, . . ., xjp) be two objects described by p numeric attributes. The Euclidean distance between objects i and j is

defined as

$$d(i,j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \cdots + (x_{ip} - x_{jp})^2}.$$

We compare irregular shape of mole to a regular ellipse created with same X and Y axis as the mole highest width and height.

4.3.2 Dice-Sorenson Co-efficient

The Sørensen–Dice coefficient (see below for other names) is a statistic used to gauge the similarity of two samples.

Sørensen's original formula was intended to be applied to discrete data. Given two sets, X and Y, it is defined as:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

where |X| and |Y| are the cardinalities of the two sets (i.e. the number of elements in each set). The Sørensen index equals twice the number of elements common to both sets divided by the sum of the number of elements in each set.

When applied to boolean data, using the definition of true positive (TP), false positive (FP), and false negative (FN), it can be written as

$$DSC = \frac{2TP}{2TP + FP + FN}.$$

The Sørensen–Dice coefficient is useful for ecological community data. Justification for its use is primarily empirical rather than theoretical (although it can be justified theoretically as the intersection of two fuzzy sets). As compared to Euclidean distance, the Sørensen distance retains sensitivity in more heterogeneous data sets and gives less weight to outliers. Recently the Dice score (and its variations, e.g. logDice taking a logarithm of it) has become popular in computer lexicography for measuring the lexical association score of two given words. It is also commonly used in image segmentation, in particular for comparing algorithm output against reference masks in medical applications.

4.4 Histogram Analysis

A histogram is an accurate representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable (CORAL) and was first introduced by Karl Pearson. It differs from a bar graph, in the sense that a bar graph relates two variables, but a histogram relates only one. To construct a histogram, the first step is to "bin" (or "bucket") the range of values—that is, divide the entire range of values into a series of intervals—and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. The bins (intervals) must be adjacent, and are often (but are not required to be) of equal size.

If the bins are of equal size, a rectangle is erected over the bin with height proportional to the frequency—the number of cases in each bin. A histogram may also be normalized to display "relative" frequencies. It then shows the proportion of cases that fall into each of several categories, with the sum of the heights equaling 1.

In image processing and photography, a color histogram is a representation of the distribution of colors in an image. For digital images, a color histogram represents the number of pixels that have colors in each of a fixed list of color ranges, that span the image's color space, the set of all possible colors.

A color histogram focuses only on the proportion of the number of different types of colors, regardless of the spatial location of the colors. The values of a color histogram are from statistics. They show the statistical distribution of colors and the essential tone of an image.

In general, as the color distributions of the foreground and background in an image are different, there might be a bimodal distribution in the histogram.

For the luminance histogram alone, there is no perfect histogram and in general, the histogram can tell whether it is over exposure or not, but there are times when you might think the image is over exposed by viewing the histogram; however, in reality it is not.

CHAPTER 5

POST-PROCESSING

5.1 Determining the Sub-Category of Skin Cancer

Now, we have 3 parameters for border irregularity and shape asymmetry: 2 Dice-Sorenson co-efficients and 1 Euclidean distance. And we have information about mole color. It is time to integrate these parameters to determine the type of skin cancer out of Malignant Melanoma, Basal Cell Carcinoma, Squamous Cell Carcinoma or Benign.

We employ “Distance finding for mixed type of attribute” technique of “Measuring Dissimilarity between two tuples”. For this, we have given equal weights to all four parameters. And some heuristics about our already acquired knowledge of skin cancer. Eg. Ulcer being bright red must lean toward Squamous Cell Carcinoma etc.

CHAPTER 6

SUPERVISED LEARNING WITH NEURAL NETWORK

6.1 Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

A Jupyter kernel to work with Python code in Jupyter notebooks and other interactive frontends.

The enhanced interactive Python shells and kernel have the following main features:

- Comprehensive object introspection.
- Input history, persistent across sessions.
- Caching of output results during a session with automatically generated references.
- Extensible tab completion, with support by default for completion of python variables and keywords, filenames and function keywords.

- Extensible system of ‘magic’ commands for controlling the environment and performing many tasks related to IPython or the operating system.
- A rich configuration system with easy switching between different setups (simpler than changing $PYTHONSTARTUP environment variables every time).
- Session logging and reloading.
- Extensible syntax processing for special purpose situations.
- Access to the system shell with user-extensible alias system.
- Easily embeddable in other Python programs and GUIs.
- Integrated access to the pdb debugger and the Python profiler.

6.2 MobileNet

Depthwise Separable Convolution is used to reduce the model size and complexity. It is particularly useful for mobile and embedded vision applications.

- Smaller model size: Fewer number of parameters
- Smaller complexity: Fewer Multiplications and Additions (Mult-Adds)

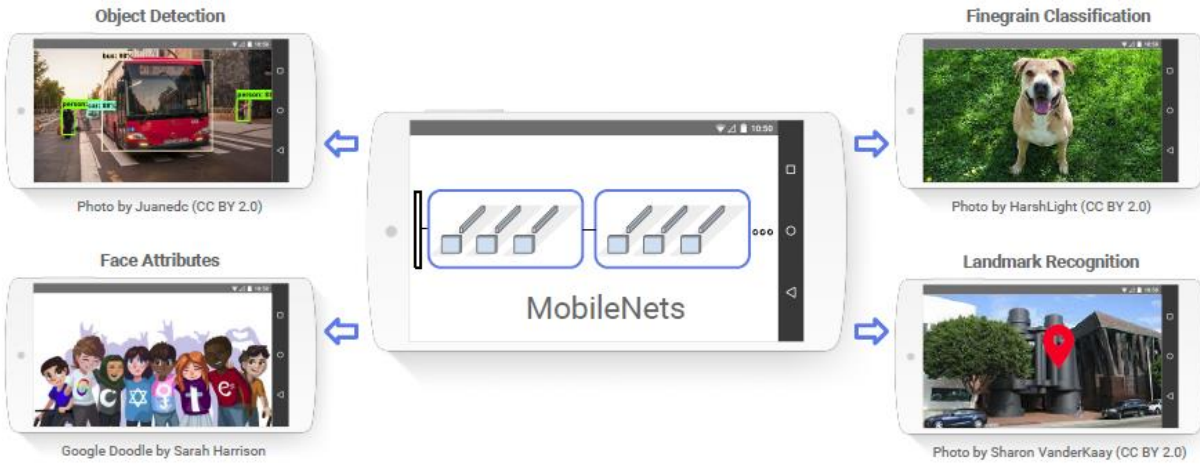


Fig 12: Applications of MobileNets

Two parameters are introduced so that MobileNet can be tuned easily: Width Multiplier α and Resolution Multiplier ρ.

6.2.1 Depthwise Separable Convolution

Depthwise separable convolution is a depthwise convolution followed by a pointwise convolution as follows:

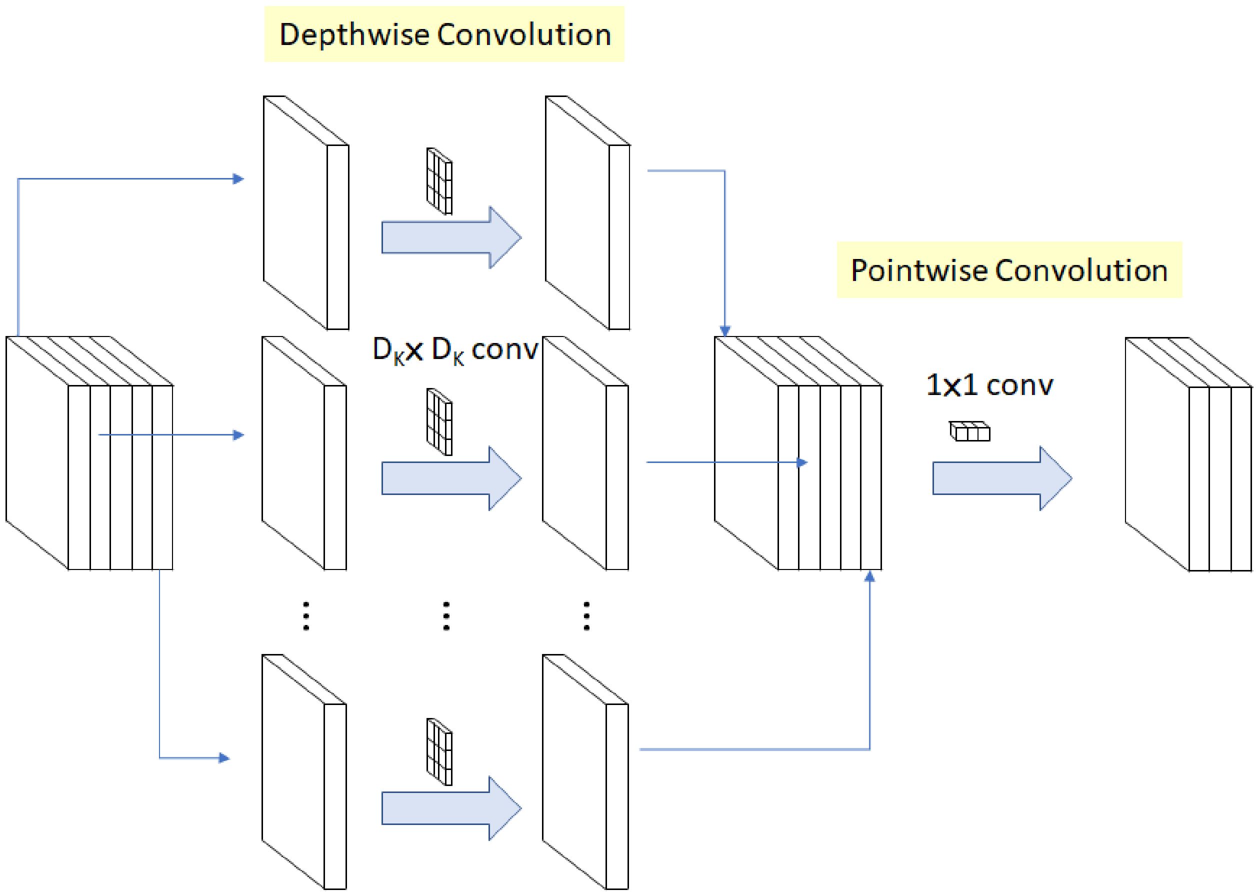


Fig 13: Architecture of MobileNets

1. Depthwise convolution is the channel-wise DK×DK spatial convolution. Suppose in the figure above, we have 5 channels, then we will have 5 DK×DK spatial convolution.

2. Pointwise convolution actually is the 1×1 convolution to change the dimension.

6.2.2 Whole Network Architecture

Below is the MobileNet Architecture:

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
5× Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
5× Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

Table 1: Layer Details of Mobilenet

6.2.3 Width Multiplier α for Thinner Models

Width Multiplier α is introduced to control the input width of a layer, which makes M become αM. Depthwise Separable Convolution Cost with Width Multiplier α

where α is between 0 to 1, with typical settings of 1, 0.75, 0.5 and 0.25. When α=1, it is the baseline MobileNet. And the computational cost and the number of parameters can be reduced quadratically by roughly α^2. Accuracy drops off smoothly from α=1 to 0.5 until α=0.25 which is too small.

6.2.4 Resolution Multiplier ρ for Reduced Representation

Resolution Multiplier ρ is introduced to control the input image resolution of the network, with Resolution Multiplier ρ .

6.2.5 Comparison With State-of-the-art Approaches

When 1.0 MobileNet-224 is used, it outperforms GoogLeNet (Winner of ILSVRC 2014) and VGGNet (1st Runner Up of ILSVRC 2014) while the multi-adds and parameters are much fewer:

Model	ImageNet Accuracy	Million Mult-Adds	Million Parameters
1.0 MobileNet-224	70.6%	569	4.2
GoogleNet	69.8%	1550	6.8
VGG 16	71.5%	15300	138

Table 2: Mobilenet compared to GoogleNet and VGG 16

To conclude, similar performance with state-of-the-art approaches but with much smaller network is achieved using MobileNet, favored by Depthwise Separable Convolution.

6.3 Confusion Matrix

In the field of machine learning and specifically the problem of statistical classification, a confusion matrix, also known as an error matrix.

A confusion matrix is a table that is often used to describe the performance of a classification model (or “classifier”) on a set of test data for which the true values are known. It allows the visualization of the performance of an algorithm.

It allows easy identification of confusion between classes e.g. one class is commonly mislabeled as the other. Most performance measures are computed from the confusion matrix.

A confusion matrix is a summary of prediction results on a classification problem.

The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix.

The confusion matrix shows the ways in which your classification model is confused when it makes predictions.

It gives us insight not only into the errors being made by a classifier but more importantly the types of errors that are being made.

	Class 1 Predicted	Class 2 Predicted
Class 1 Actual	TP	FN
Class 2 Actual	FP	TN

Table 3: Confusion Matrix Skeleton

CHAPTER 7

EXPERIMENTS AND RESULTS

7.1 Image Processing Throughput

Clearly, the first thing image processing efficiency depends on is dimensions of image.

The following table shows some of the results for different sizes of image:

Serial Number	Image Size (width * height)	Total time taken for predictions
1	600 pixel x 450 pixel	14.945 s
2		
3	2048 pixel x 1536 pixel	37.448 s
4	6668 pixel x 4419 pixel	2737.407 s

Table 4: Performance analysis

Profile Summary

Generated 22-Apr-2019 09:47:42 using performance time.

Function Name	Calls	Total Time	Self Time*	Total Time Plot (dark band = self time)
WholeProcess	1	32.541 s	18.570 s	

Fig 14: Time taken to process one 2112 x 2816 image

7.2 Neural Network Throughput

The idea of using training data in machine learning programs is a simple concept, but it is also very foundational to the way that these technologies work. The training data is an initial set of data used to help a program understand how to apply technologies like neural networks to learn and produce sophisticated results. It may be complemented by subsequent sets of data called validation and testing sets.

The training set is the material through which the computer learns how to process information. Machine learning uses algorithms – it mimics the abilities of the human brain to take in diverse inputs and weigh them, in order to produce activations in the brain, in the individual neurons. Artificial neurons replicate a lot of this process with software – machine learning and neural network programs that provide highly detailed models of how our human thought processes work.

With that in mind, training data can be structured in different ways. For sequential decision trees and those types of algorithms, it would be a set of raw text or alphanumerical data that gets classified or otherwise manipulated. On the other hand, for convolutional neural networks that have to do with image processing and computer vision, the training set is often composed of large numbers of images. The idea is that because the machine learning program is so complex and so sophisticated, it uses iterative training on each of those images to eventually be able to recognize features, shapes and even subjects such as people or animals. The training data is absolutely essential to the process – it can be thought of as the "food" the system uses to operate.

```
Confusion matrix, without normalization
[[1816    4]
 [  11 1809]]
```

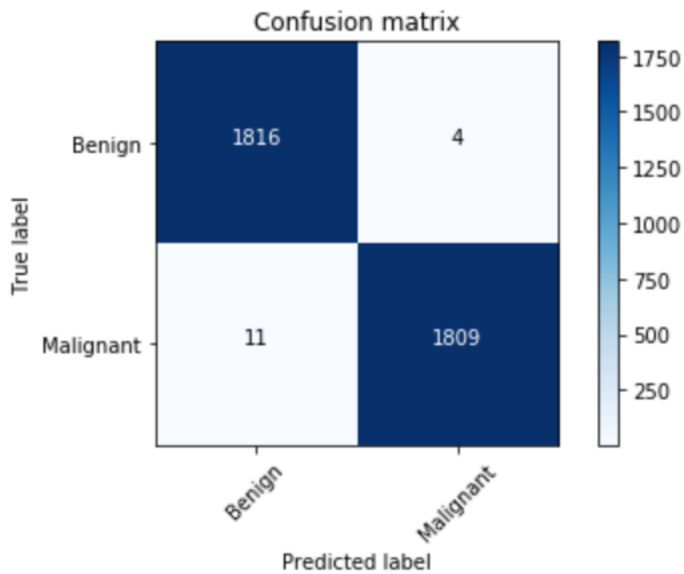


Fig 15: Confusion Matrix for Training dataset prediction

Train Predictions: 3640/3640 [==============================] - 1363s

When using the trained data as test data after the training phase has been completed, we get the above results. It takes 1363 seconds to run predictions for 3640 images.

In order to test a software application, you need to enter some data for testing most of the features. Any such specifically identified data which is used in tests is known as test data.

You can have test data in excel sheet which can be entered manually while executing test cases or it can be read automatically from files (XML, Flat Files, Database etc.) by automation tools.

Some test data is used to confirm the expected result, i.e. When test data is entered the expected result should come and some test data is used to verify the software behavior to invalid input data.

Test data is generated by testers or by automation tools which support testing. Most of the times in regression testing the test data is re-used, it is always a good practice to verify the test data before re-using it in any kind of test.

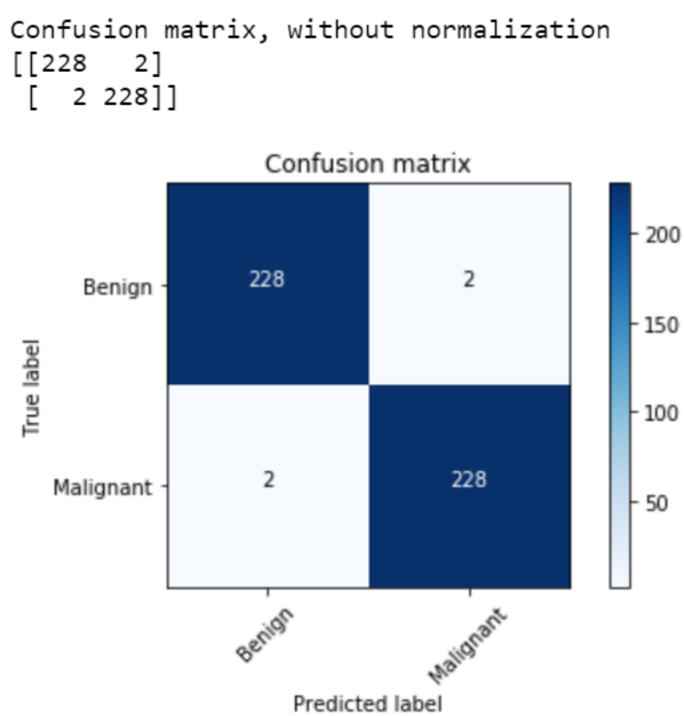


Fig 16: Confusion Matrix for Testing dataset prediction

Test Predictions: 460/460 [==============================] - 139s

When using the test dataset after the training phase has been completed, we get the above results. It takes 139 seconds to run predictions for 460 images.

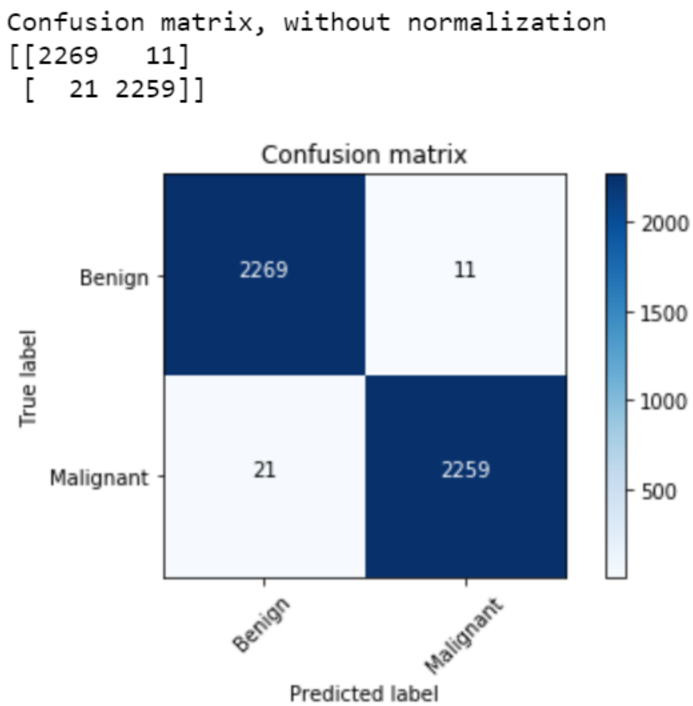


Fig 17: Confusion Matrix for all dataset prediction

All Predictions: 4560/4560 [==============================] - 1364s

The results after mixing all data together. To get a thorough idea of prediction model accuracy.

CHAPTER 8

CONCLUSION AND FUTURE WORK

- The neural network had only enough data for Malignant Melanoma and Benign mole to work on predictions. Finding more data for Basal and Squamous Cell Carcinoma may result in some changes in current prediction success rate, so we must take that into account before reaching a conclusion of comparison between unsupervised and supervised learning.
- A better measure than just using heuristics for image processing needs to be found. But for now, because of the lack of equal amount of data for all 4 classes, heuristics were employed.
- Many dermatoscope images were captured with a round apparatus which results in images having a black artefact on the border. Most obvious path for now is to manually crop that part because it's hard to predict how much of the black color is because of the apparatus and how much is of the mole (ROI).
- Hair and Scale mark detection system can miss fine hairs or very thin scale marks but as they are very fine in granularity anyway, that doesn't change or skew the results drastically. Still, more work can be done in this area.

REFERENCES

[1] Artefact Removal and Contrast Enhancement for Dermoscopic Images Using Image Processing Techniques; Pragati Rajendra Mahajan1, Prof. Mrs. A. J. Vyavahare; Dec 9, 2013; ISSN (Online) 2321 – 2004; International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering Vol 1, Issue 9

[2] Cancer Cells Detection Using Digital Image Processing Methods; Bhagyashri G. Patil, Prof. Sanjeev N. Jain; March 2014; ISSN: 2278-621X; International Journal of Latest Trends in Engineering and Technology (IJLTET) Vol. 3 Issue 4

[3] Computer aided Melanoma skin cancer detection using Image Processing; Shivangi Jaina, Vandana Jagtapb, Nitin Pise; (ICCC-2015); International Conference on Intelligent Computing, Communication & Convergence

[4] Diagnosis of Skin Cancer Using Image Processing; Esperanza Guerra-Rosasa, Josué Álvarez-Borregob, Ángel Coronel-Beltrán; Oct 6, 2014; International Conference of Computational Methods in Sciences and Engineering; doi: 10.1063/1.4897704

[5] Segmentation of Skin Cancer Images; Padmapriya Nammalwar, Ovidiu Ghita, Paul F. Whelan

[6] Skin cancer detection and stage prediction using image processing techniques; Sheeju Diana B, Ramamurthy B; International Journal of Engineering & Technology; doi: 10.14419/ijet.v7i1.8643

[7] https://www.cancer.org/cancer/skin-cancer/prevention-and-early-detection.html

[8] https://www.isic-archive.com/#!/topWithHeader/onlyHeaderTop/gallery

[9] https://link.springer.com/chapter/10.1007%2F978-0-387-77574-6_8

APPENDIX

Screenshots:

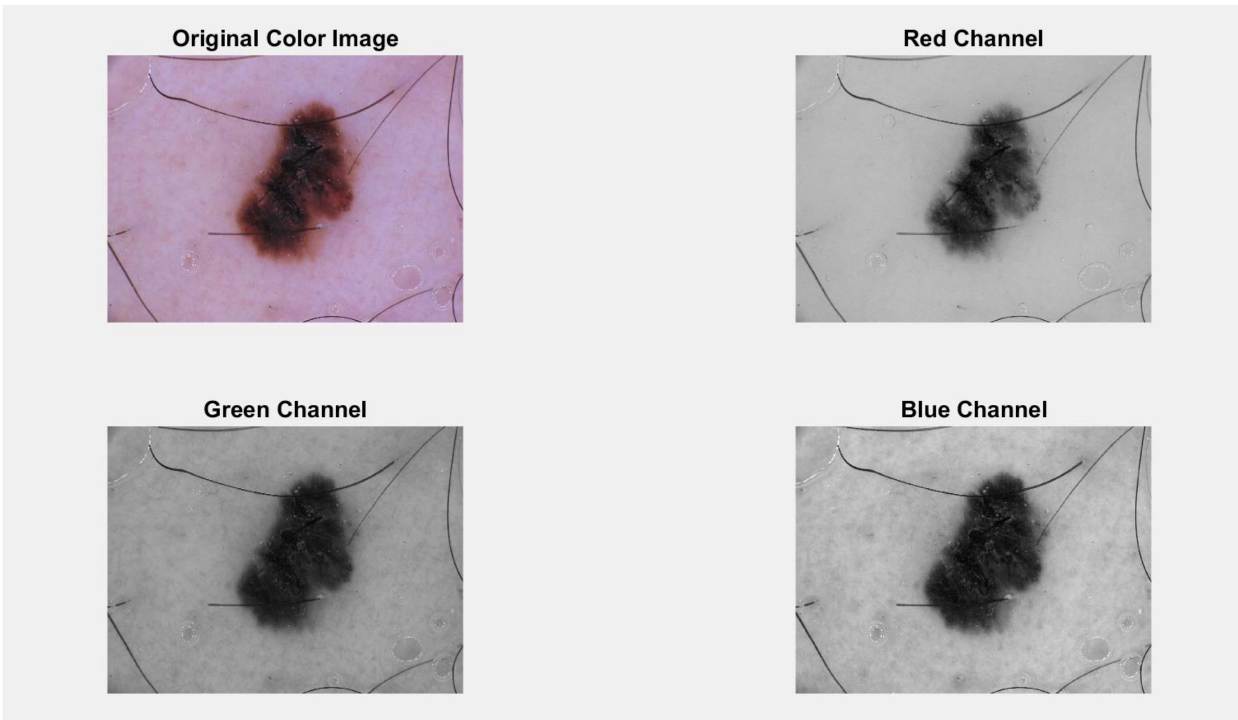


Fig 18: Extraction of RBG Channels from color image

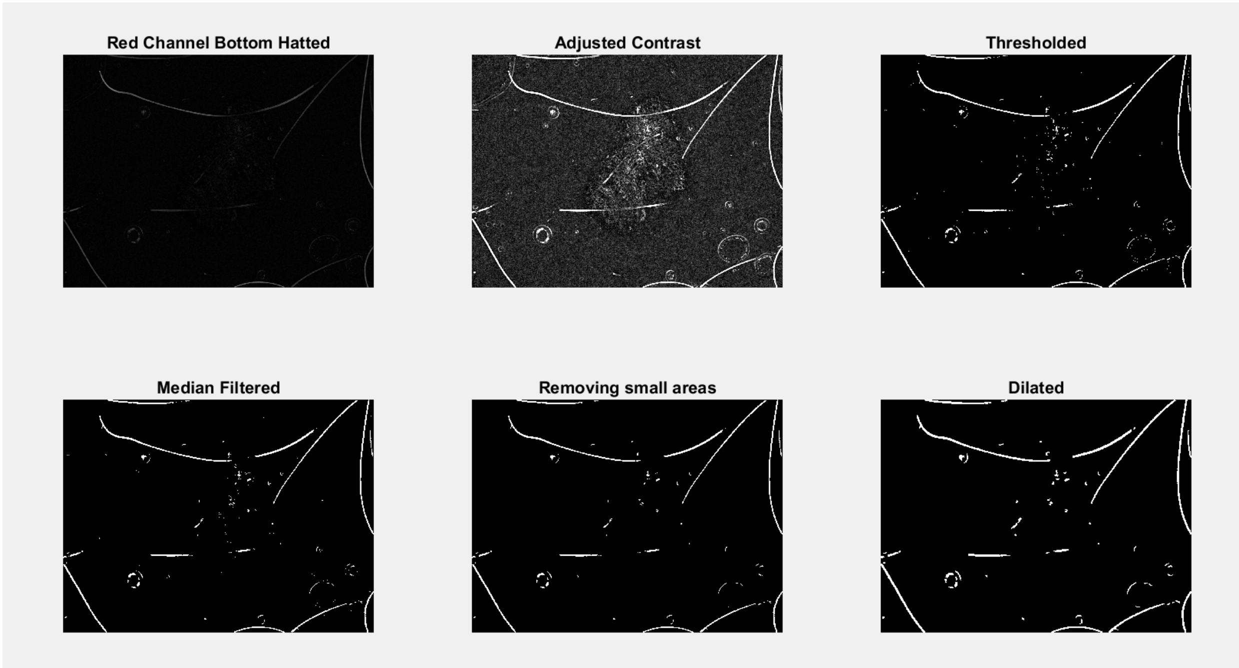


Fig 19: Extraction of Hair and/ or Scale marks from Red channel

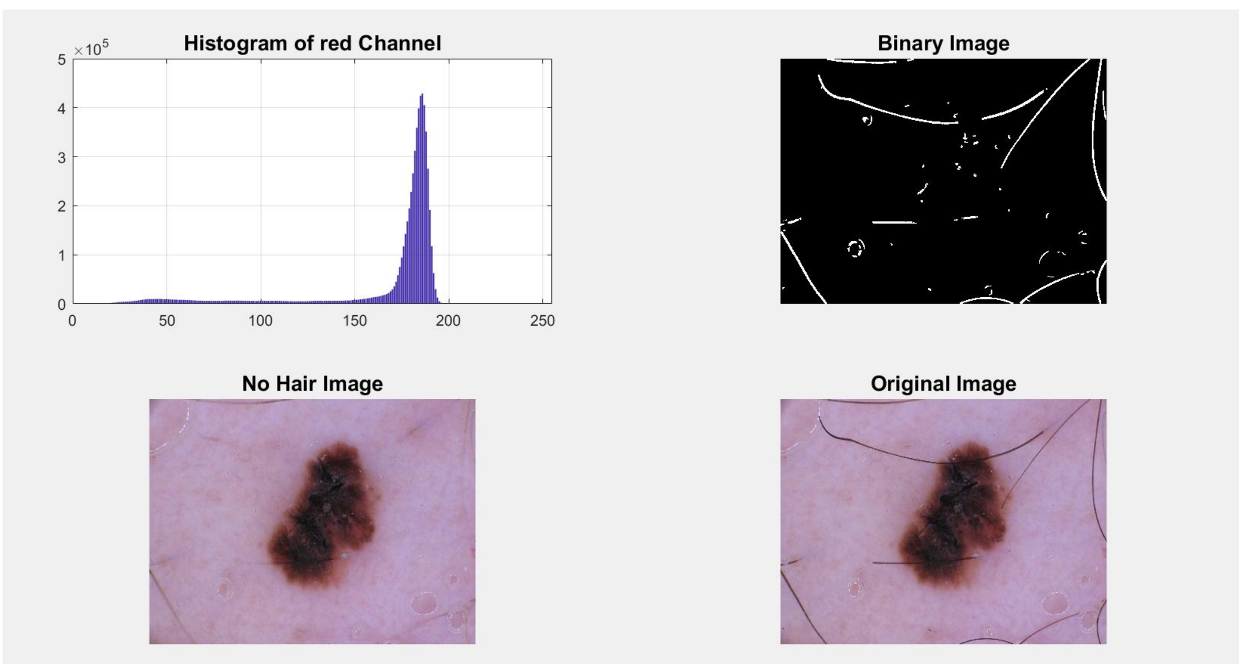


Fig 20: Region filling Hair and/ or Scale marks in original RGB image

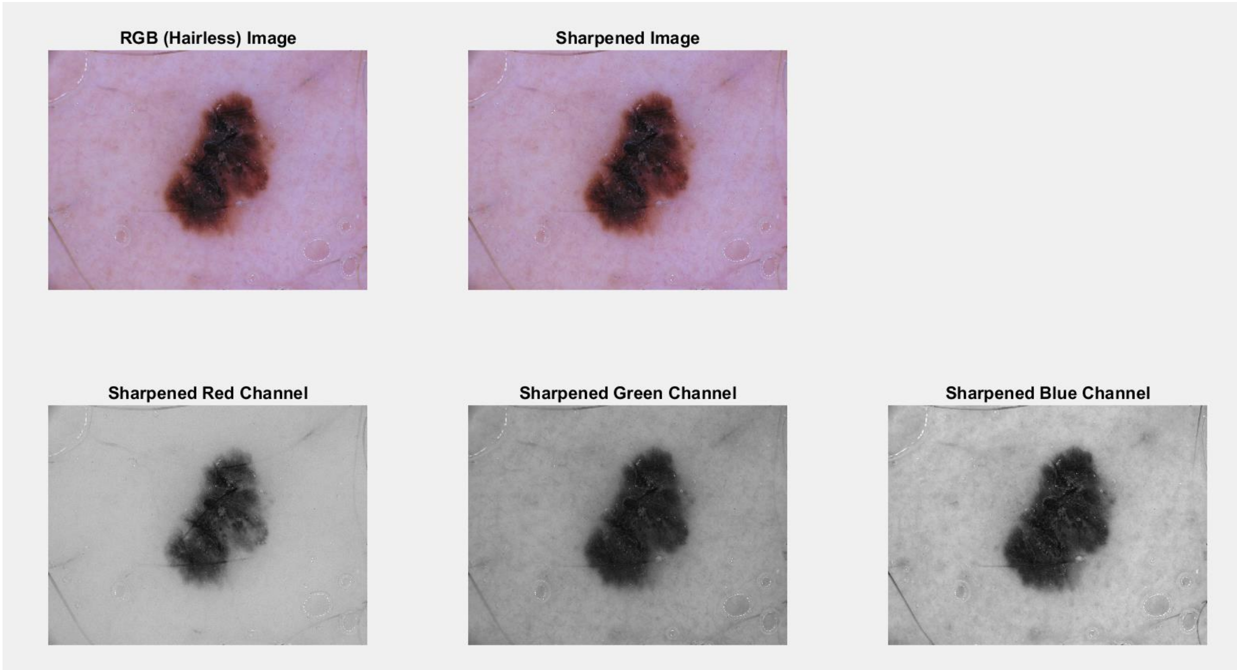


Fig 21: Adjusting the contrast of hairless RGB image

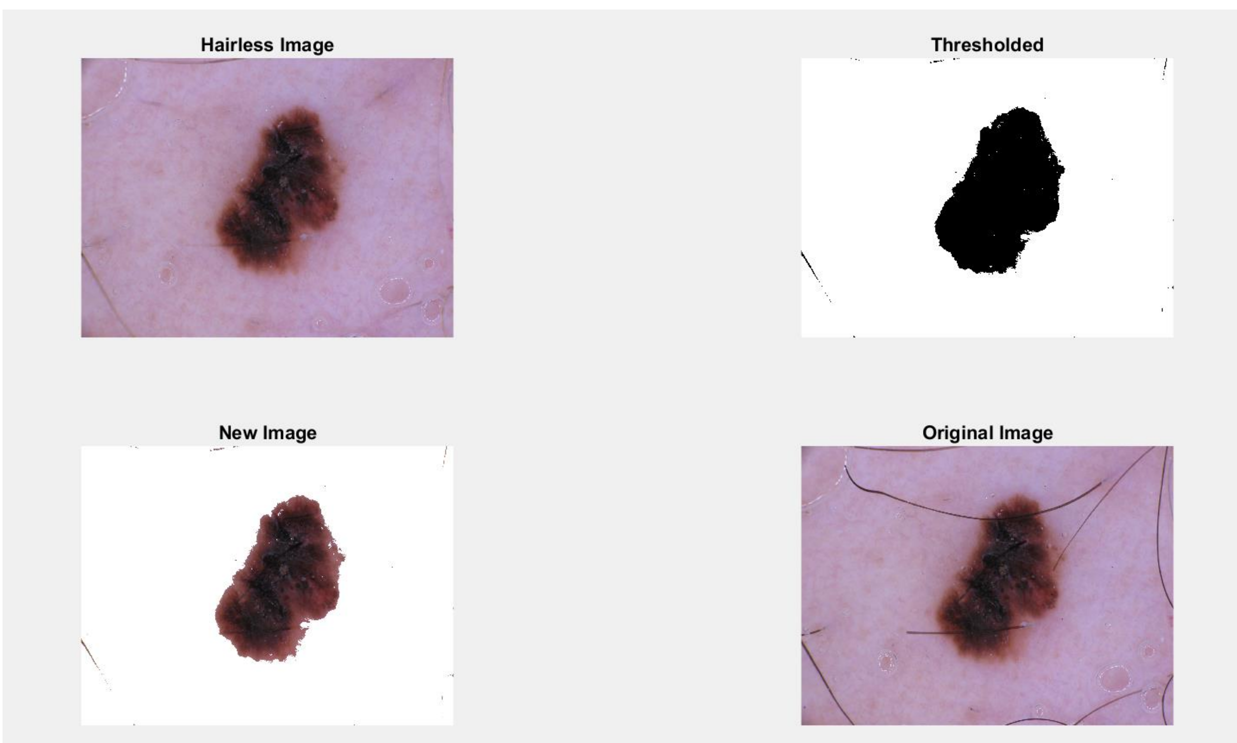


Fig 22: Detecting Region Of Interest from hairless RGB image

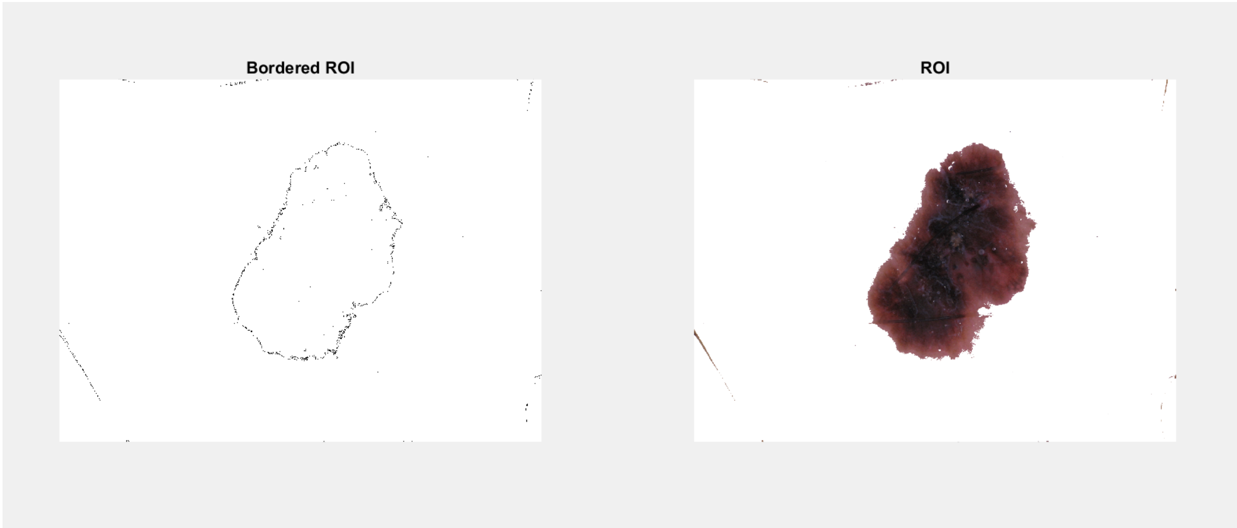


Fig 23: Detecting borders of Region Of Interest

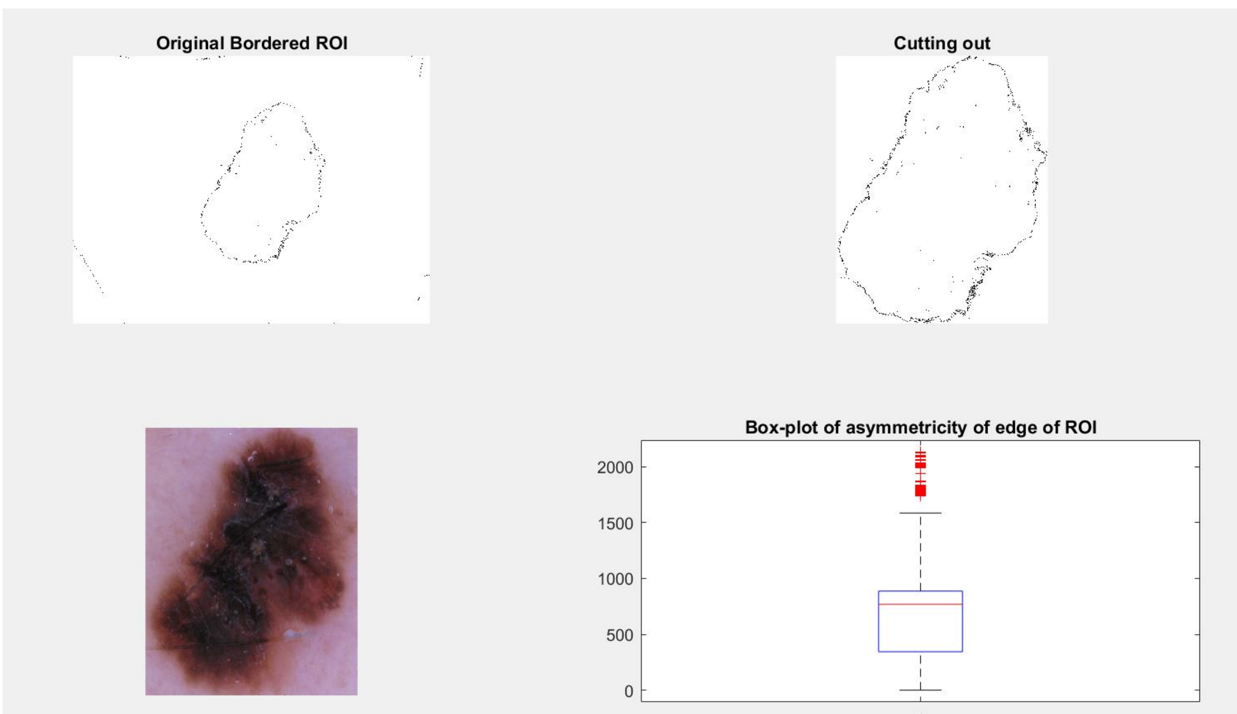


Fig 24: Cutting out Region Of Interest and plotting asymmtricity box-lot which is calculated using Euclidean distance.

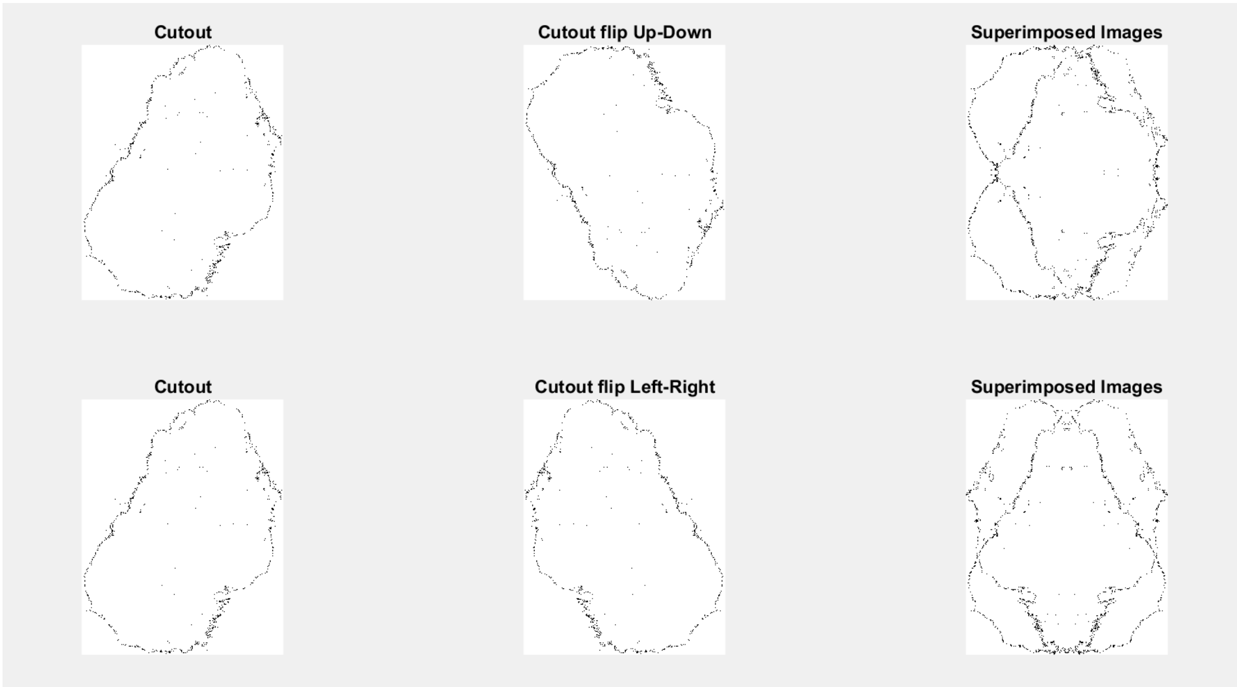


Fig 25: Finding the value of Dice-Sorenson Co-efficient

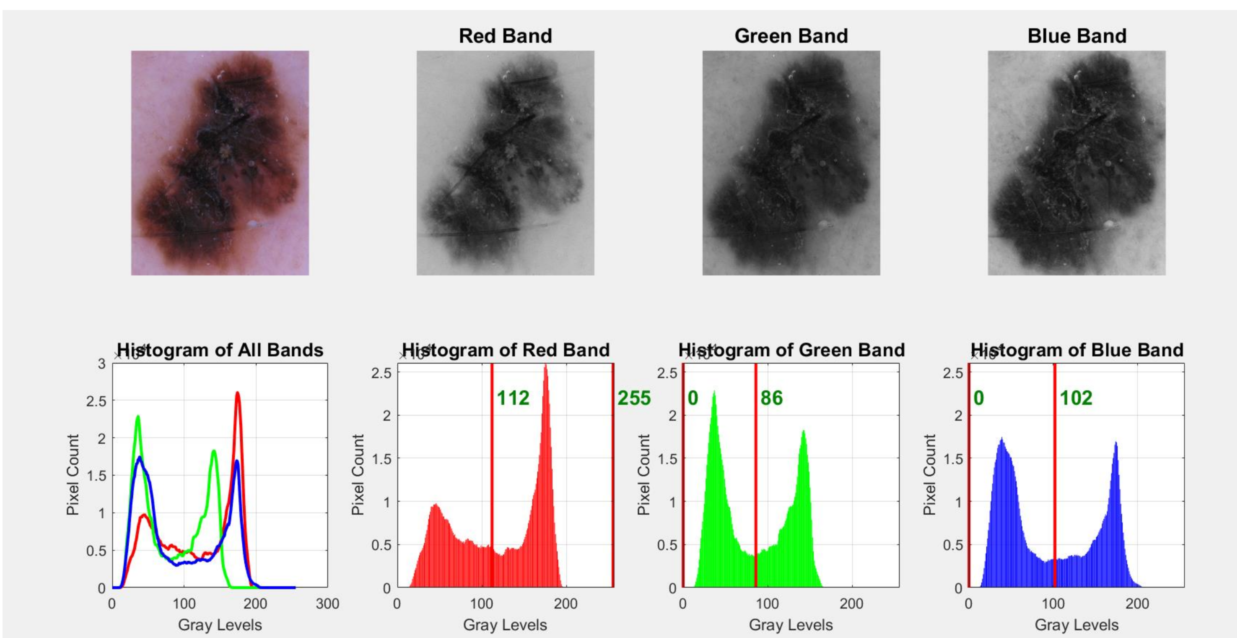


Fig 26: Finding the color of mole/ ROI by Histogram Analysis